Maternal & Child Nutrition

DOI: 10.1111/j.1740-8709.2009.00221.x

Original Article

Reproducibility and validity of a food frequency questionnaire in assessing dietary intakes of low-income Caucasian postpartum women living in Sheffield, United Kingdom

Theodora Mouratidou, Fiona A. Ford and Robert B. Fraser

Academic Unit of Reproductive and Developmental Medicine, University of Sheffield, the Jessop Wing Royal Hallamshire Hospital, Tree Root Walk, Sheffield S10 2SF, UK

Abstract

The aim of this study was to examine the reproducibility and validity of a semi-quantitative food frequency questionnaire (FFQ) for assessing dietary intakes of low-income, Caucasian, English-speaking, postpartum women living in Sheffield, United Kingdom. Data was obtained from a cross-sectional sample of the 'Healthy Start' study; a population-based survey of mothers and infants. Participants completed two FFQs at 4 and 8 weeks postpartum. Measures from 24-hour dietary recalls (24HDRs) were collected at 4, 6, 8 and 12 weeks postpartum. In the reproducibility study, crude Pearson's correlation coefficients ranged from 0.40 (riboflavin) to 0.73 (thiamine), mean value 0.54. In the validation study, crude Pearson correlation coefficients between the FFQ and the measures from the 24HDRs ranged from 0.10 (B12) to 0.55 (manganese), mean value 0.34. Energy-adjustments and corrections for attenuation had no significant effect on the strength of the correlation both observed in the reproducibility and validity study. On average, 68% of the participants were classified correctly, and 3% were misclassified into the extreme opposite quintile of the distribution. The authors conclude that the questionnaire performed well for the majority of nutrients examined and that is a valid tool for ranking individuals according to nutrient distribution.

Keywords: food frequency questionnaire, postpartum nutrition, reproducibility, validity.

Correspondence: Theodora Mouratidou, Academic Unit of Reproductive and Developmental Medicine, University of Sheffield, the Jessop Wing Royal Hallamshire Hospital, Tree Root Walk, Sheffield S10 2SF, UK. E-mail: t.mouratidou@sheffield.ac.uk

Introduction

A series of global and country-specific policy initiatives are in place for the promotion of maternal and infant health. The World Health Organization's European Second Action Plan was developed to address the main public health challenges in the field of nutrition. The plan indicates the promotion of optimal fetal, infant and young children's nutrition as a major action area in the challenge of tackling diet-related non-communicable diseases (World Health Organization, Regional Office for Europe 2007).

Maternal personal food preferences and dietary decisions for their family play an important part in shaping the diet of their infants and young children (Birch & Fisher 1998; Jain *et al*. 2001). Even though the health of mothers is of particular importance for theirs and their children's future health as well as for any future pregnancies, limited data on the nutritional status of women in the first year of a child's life is available. Related nutritional concerns include dietary inadequacies, iron deficiency anaemia, low calcium intakes, and low or excess pregnancy weight gain (Doran & Evers 1997; Mackey *et al*. 1998;

Rees *et al*. 2005). Pregnancy and postpartum are significant life transition periods for lactating and non-lactating women and might influence dietary behaviours and diet-related practices of some women (George *et al*. 2005).

Despite the less accurate approximation of dietary intakes of individuals compared with other methods, the food frequency method is often used to assess dietary intakes in a wide variety of settings and populations mainly because of its cost effectiveness. The dietary recall method is often used as reference method for validating FFQs because it is a less demanding method and less likely to influence the actual diet of the participants than other dietary methods e.g. weighted records (Willett 1998; Gibney *et al*. 2004).

The objective of this study was to evaluate the reproducibility and validity of an interviewer-administered FFQ in a population of low-income, Caucasian (white-British) women living in Sheffield, UK in the first weeks after delivery.

Materials and methods

Study design and sample

Participants were drawn from the 'Healthy Start' (HS) project in Sheffield, UK, which is a cohort study investigating maternal and infant nutrition practices before (phase 1) and after (phase 2) the introduction of a new UK government food-benefit scheme called HS (Ford *et al*. 2009). Under the HS scheme low-income families are entitled to vitamin supplements and vouchers which can be exchange for fresh fruit and vegetables as well as milk and infant formula milk (Department of Health 2002).

Initial recruitment of all HS study participants took place at the postnatal wards of the Jessop Wing, Royal Hallamshire Hospital in Sheffield, UK where an explanatory information leaflet was given to them. They were then asked for their permission to be telephoned by a member of the research group once returned home, to ask whether they would like to participate in the study or not. Face-to-face, interviewer-administered questionnaire interviews were carried out at 4 weeks postpartum by trained research assistants and thereafter over the phone at each month during the first year of the baby's life.

Participants were included in the study only if they were of Caucasian ethnic origin (white British), English speaking, living in Sheffield, free of any nutrition-related pre-existing medical condition such as diabetes or coeliac disease, low socio-economic status, and had a live, healthy baby. Postcodes were used to identify subjects living in deprived electoral wards of Sheffield using the Index of Multiple Deprivation 2004 (Noble *et al*. 2004). Therefore, area of residence was used as a proxy of low-income status. Data used in this analysis were collected between November 2005 and November 2006 (phase 1), and dietary intakes were obtained via multiple FFQs and 24HDRs (Fig. 1).

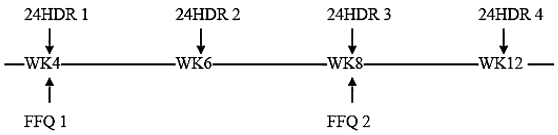

Fig. 1. The design of the validation study conducted among 104 Caucasi postpartum woman, November 2005 to November 2006. FFQ, food frequency questionnaire; WK, week; 24HDR, 24-hour dietary recall.

Key messages

- Maternal personal food preferences and dietary decisions for their family play an important part in shaping the diet of their infants and young children. Therefore, more research is needed on maternal food choices, dietary intakes and eating habits.
- The study findings suggested good reproducibility and fair validity for the majority of nutrients examined as assessed by the interviewer-administered FFQ and multiple 24HDRs.
- The FFQ was able to distinguish between high and low consumers for all nutrients under investigation. Hence, it is a valid tool for ranking individuals according to nutrient distribution even if absolute intakes may not be that precise.

One hundred and forty-one participants fulfilled the specified criteria and were recruited. Thirty seven women were removed from the dataset because of incomplete follow-up dietary data (either FFQs or 24HDRs). In total, 104 women were included in the validity analysis (FFQ1 compared with the measures from the 24HDRs), which is 50% of the total HS study participants. For the purposes of the reproducibility analysis, out of the 104, 23 participants were further removed from the dataset because of an incomplete second FFQ. In total, 81 participants were included in the reproducibility analysis (FFQ1 compared with FFQ2), which is 39% of the total HS study participants.

This study was performed as a Service Evaluation with the approval of the North Sheffield Local Research Ethics Committee.

Semi-quantitative FFQ

The interviewer-administered FFQ is an adaptation of the FFQ developed by Rogers and colleagues (Rogers *et al*. 1998) and has been validated for use in pregnancy (Mouratidou *et al*. 2006a). The FFQ includes 62 quantitative and qualitative questions, 40 of which assessed the frequencies of consumption of meat, poultry, fish and seafood, common vegetables and fruits, cereals and confectionery. There are also detailed questions about the frequency, type and amount of fat, bread, alcohol and milk consumed. Participants were asked to report frequency of food consumption over the 4 weeks prior to administration. The frequency options included: never or rarely, once a fortnight, 1–3 times a week, 4–7 times a week and more than once a day. To increase the simplicity and the inclusion of a wide range of food eaten, no portion size quantification was asked; thus standard portion sizes were assumed throughout.

24-hour dietary recalls

The reference method was a series of 24HDRs (1st recall obtained during a home visit and following recalls over the phone). Participants were aware of the time frame for when the next 24HDRs would take place but not the exact date. Participants were asked to recall all foods and drinks consumed the previous day. Household measurements were used to estimate portion sizes. At the end of the interview the foods were summarized for the respondent. During the follow-up period, efforts were made to ensure that each participant was interviewed various days of the week therefore for a number of participants weekend dietary information was randomly collected. Mean intakes from the 24HDRs were calculated by averaging the crude intakes from the 4 recalls obtained at each time point. Small random-within person variation observed in dietary intakes by time point allowed inclusion of subjects who completed only 2 or 3 recalls, and a third or fourth reading was taken from another time point when available i.e. 16 week postpartum. Twenty-two per cent of the 104 participants completed a 24HDR at 16 weeks.

Nutrient calculation

The nutritional software systems used for nutrient calculations have been described elsewhere (Mouratidou *et al*. 2006b). In brief, daily intakes of energy and nutrients obtained from the FFQs were analyzed using Q-Builder (Tinuviel Software, Anglesey, UK) which converts information on food consumption into a list of foods and weights. Approximate daily intake is then calculated by multiplying the weekly frequency of consumption of a food by the nutrient content of a standard portion. Each one of the frequency options the questionnaire allocated was mapped as follows: never or rarely = 0, once a fortnight = 0.5, 1–3 times a week = 2, 4–7 times a week = 5.5 and more than once a day = 14. The daily mean intakes obtained from the 24HDRs were analyzed using WISP (Tinuviel Software, Anglesey, UK).

Statistical methods

Statistical analyses were performed using the Statistical Package for the Social Sciences version 14.0 (SPSS Inc., Chicago, IL, USA). Daily intakes examined are presented as means from foods only. Distributions of intake for most nutrients were skewed towards higher values, therefore, crude data from FFQ1, FFQ2 and the mean of the 24HDRs were

log-transformed to improve normality before calculating means and confidence intervals (CI).

Pearson product-moment correlation coefficients were calculated to assess the reproducibility of FFQ1 compared with FFQ2. Pearson coefficients were calculated and compared with Spearman's rank coefficients for log-transformed values. The results were rather similar and only Pearson coefficients are presented. Agreement between FFQ1 and 24HDR was examined using the Bland-Altman analysis where the mean difference between the two measurements ($FFQ1_{log}$ – $24HDRs_{log}$) is plotted against their mean [($FFQ1_{log}$ + $24HDR_{log}$)/2]. The analysis assesses agreement in individuals, defined as the limit of agreements (LoA) (±2SD of the bias) (Bland & Altman 1986). To permit interpretation of the log-transformed data, the antilog of the data was calculated which represents the FFQ/24HDR ratio. Pearson coefficients were also calculated to evaluate the relative validity of FFQ1.

In addition to comparisons based on crude dietary intakes, comparisons were also made using energy-adjusted values. Adjustment for total energy intake was done using the residual method (Willett *et al*. 1997). All validity coefficients were corrected for attenuation due to random error in within-person variability in the 24HDRs. Variance component analysis was used to calculate the within- and between-person variation in the 24HDRs. The correction for the attenuating effect of random within-person error was computed according to the following equation:

$$r_{adjusted} = r_{observed}\sqrt{1+\lambda_x/n_x}$$

where λ_x is the ratio of the within- and between-person variances for x, and n_x is the number of replicates for the x variable (Willett 1998).

Therefore, reproducibility coefficients are presented as unadjusted and adjusted for total energy intake. Validation coefficients are presented as unadjusted, energy-adjusted and de-attenuated coefficients adjusted for total energy intake.

The distribution of crude and energy-adjusted nutrient intakes was also divided into quintiles and the proportion of correctly categorized subjects in the same (correct classification) or into extreme opposite quintiles (misclassification) was then calculated.

Results

Table 1 presents selected anthropometric, demographic and behavioural characteristics of the study participants. The average age (SD) of the participants was 23 (5.0) years of age, and the average body mass index (BMI) is 26 (6.0). Twenty-five per cent of the women classified themselves as not having achieved any educational qualifications and a high proportion was categorized as self-reported smokers (35%).

Reproducibility

Mean nutrient intakes estimated by the FFQ1 and FFQ2 are presented in Table 2. Mean intakes of all nutrients examined were higher for FFQ1 than FFQ2 and differences were within a range of 7%. FFQ2 gave 1 to 18% lower values than the FFQ1 with the most noticeable difference observed for saturated fatty acids (18%). Crude Pearson correlation coeffi-

Table 1. Characteristics of study participants

	Mean (SD)
Age (years)*	23 (5.0)
BMI (kg/m^2)†	26 (6.0)
	(%)
Maternal age in categories (years)*	
<19	15
20–24.9	27
25–34.9	50
>34	8
Educational attainment	
5 GCSE's or more	75
No qualification	25
Self-reported smoking status	
Non-smoker	65
Current smoker	35
Parity	
0	58
1†	42

GCSE, general certificate of secondary education. *n = 101; †n = 79.

Table 2. Reproducibility study: mean dietary intakes and Pearson correlation coefficients between mean dietary intakes based on FFQ1 and FFQ2

Nutrients	FFQ1		FFQ2		% of FFQ1	Pearson *r*	Pearson *r* adjusted†
	FFQ1 (mean)	(95% CI)	FFQ2 (mean)	(95% CI)			
Energy (MJ)	8.1	4.8, 13.8	7.3	4.5, 11.7	90	0.54*	
Protein (g)	69	40, 116	63	35, 113	92	0.56*	0.60*
Total fat (g)	84	45, 157	71	39, 128	85	0.53*	–0.21**
Carbohydrate (g)	239	136, 418	222	137, 358	93	0.56*	0.45*
Saturated fatty acids (g)	32	16, 65	27	14, 50	82	0.53*	0.45*
Monounsaturated fatty acids (g)	27	14, 52	23	12, 44	85	0.56*	0.53*
Polyunsaturated fatty acids (g)	14	6.3, 29	12	6.0, 25	91	0.64*	0.60*
Sugars (g)	91	38, 219	86	40, 183	94	0.57*	0.57*
Starch (g)	138	81, 233	127	73, 220	92	0.53*	0.44*
Dietary fibre (g)	14	7.3, 25	13	7.0, 26	97	0.65*	0.57*
Calcium (mg)	801	396, 1620	685	369, 1273	86	0.49*	0.59*
Magnesium (mg)	252	147, 435	240	137, 418	95	0.55*	0.41*
Iron (mg)	11	6.4, 19	10.71	5.6, 20	97	0.54*	0.61*
Zinc (mg)	8.4	4.9, 14	7.8	4.6, 13	94	0.52*	0.67*
Selenium (μg)	43	21, 89	41	22, 78	95	0.46*	0.48*
Vitamin D (μg)	3.9	1.9, 8.1	3.8	1.9, 7.4	96	0.59*	0.58*
Vitamin E (mg)	5.2	3.0, 9.0	4.9	2.8, 8.6	94	0.47*	0.56*
Thiamine (mg)	2.6	1.2, 5.6	2.5	1.3, 4.7	95	0.73*	0.58*
Riboflavin (mg)	2.5	1.7, 3.6	2.4	1.6, 3.5	95	0.40*	0.63*
Vitamin B6 (mg)	3.0	2.1, 4.2	2.9	2.0, 4.3	98	0.49*	0.57*
Vitamin B12 (μg)	4.6	2.3, 9.3	4.5	2.3, 9.0	99	0.45*	0.59*
Folate (μg)	235	129, 426	224	122, 411	95	0.48*	0.58*
Vitamin C (mg)	65	25, 171	64	26, 156	98	0.60*	0.51*

*$P < 0.01$(2-tailed); **$P > 0.05$ (2-tailed). †Adjusted for energy intake.

cients ranged from 0.40 for riboflavin to 0.73 for thiamine with a mean of 0.54. Coefficients for energy-adjusted nutrient estimates varied from –0.21 for total fat to 0.67 for zinc. Average energy-adjusted coefficient was 0.52. Energy-adjustments decreased the correlation coefficient values of all macronutrients examined with the exception of protein. Adjustments slightly increased most vitamin and mineral values.

Validity

Nutrient intake estimates were higher for FFQ1 than the mean of the 24HDRs (Table 3). Pearson correlation coefficients ranged from 0.10 for vitamin B_{12} to 0.55 for manganese. The average crude coefficient for all nutrients was 0.34. For the majority of the nutrients, correlations reached statistical significance at $P < 0.01$ level except for total fat, selenium and vitamin D whose correlations were significant at $P < 0.05$ level. Correlation for vitamin B_{12} was not significant.

After adjustments for total energy intake coefficient values remained fairly modest and decreased for the majority of the nutrients. Values varied from –0.01 for vitamin B_{12} to 0.58 for manganese. Exceptions included iron, calcium, vitamin E, thiamine and vitamin C where a noticeable increase in the coefficients was observed i.e. unadjusted value for iron was $r = 0.49$ and adjusted $r = 0.57$. Coefficients for protein and total fat remained similar but coefficient for carbohydrate was attenuated i.e. 0.30 to 0.15. The average energy-adjusted coefficient for nutrients was 0.32. When corrected for the effect of random within-person variation, energy-adjusted and de-attenuated correlations ranged from –0.01 for vitamin B_{12} to 0.60 for manganese. Modest improvement in the energy-adjusted coefficient values were observed for 11 out of the 23 nutrients examined. Average de-attenuated coefficient was 0.33.

The extent to which the two methods agreed for individual and group mean intakes was also examined using Bland-Altman analysis. Figure 2 illustrates iron

Table 3. Validation study: mean nutrients intakes and Pearson correlation coefficients between mean nutrient intakes based on FFQs and 24HDRs and Bland-Altman analysis

Nutrient	FFQ (mean)	(95%CI)	24HR (mean)	(95%CI)	Pearson correlation coefficients between FFQ and recalls			Bland-Altman	
					Unadjusted	Energy adjusted	De-attenuated and energy adjusted	Mean† FFQ/ 24HDR ratio	95% Limits of agreement
Energy (MJ)	8.1	4.8, 13	5.8	4.3, 7.7	0.32*			1.4	0.74–2.6
Protein (g)	68	39, 120	53	39, 73	0.34*	0.31*	0.33*	1.2	0.64–2.5
Total fat (g)	83	46, 151	55	39, 78	0.23**	0.23**	0.24*	1.5	0.68–3.3
Carbohydrate (g)	242	139, 420	179	135, 237	0.30*	0.15***	0.16*	1.3	0.70–2.6
Saturated fatty acids (g)	32	16, 62	21	15, 31	0.28*	0.24*	0.26*	1.4	0.65–3.4
Monounsaturated fatty acids (g)	27	14, 50	18	12, 26	0.25*	0.24**	0.26*	1.5	0.65–3.4
Polyunsaturated fatty acids (g)	14	6.7, 28	9.1	6.0, 14	0.33*	0.32**	0.34*	1.5	0.62–3.6
Sugars(g)	93	39, 221	69	42, 111	0.34*	0.30*	0.31*	1.3	0.47–3.8
Starch (g)	139	83, 233	99	72, 135	0.26*	0.14***	0.15*	1.4	0.70–2.8
Dietary fibre (g)	13	7.0, 26	9	6.3, 13	0.34*	0.30*	0.33*	1.4	0.67–3.3
Calcium (mg)	778	390, 1549	551	380, 798	0.29*	0.24**	0.26*	1.4	0.60–3.3
Magnesium (mg)	249	138, 448	167	123, 226	0.45*	0.50*	0.52*	1.4	0.80–2.7
Iron (mg)	11	6, 20	7.9	5.9, 11	0.49*	0.57*	0.59*	1.3	0.76–2.5
Zinc (mg)	8.3	4.6, 14	6.9	5.1, 9.3	0.40*	0.33*	0.37*	1.1	0.63–2.2
Manganese (mg)	3.5	2.1, 6.1	2.8	2.2, 3.7	0.55*	0.58*	0.60*	1.2	0.76–2.0
Selenium (μg)	41	19, 90	25	16, 40	0.21**	0.12***	0.13*	1.6	0.56–4.8
Vitamin D (μg)	3.8	1.8, 8.2	2	1.4, 3	0.21**	0.20**	0.22*	1.8	0.73–4.7
Vitamin E (mg)	5.2	3.1, 8.7	5.5	3.7, 8.3	0.18*	0.28*	0.30*	0.94	0.39–2.2
Thiamine (mg)	2.6	1.2, 5.7	2	1.6, 2.7	0.50*	0.51*	0.54*	1.2	0.64–2.5
Riboflavin (mg)	2.4	1.7, 3.6	2	1.7, 2.6	0.34*	0.25*	0.26*	1.1	0.74–1.9
Vitamin B6 (mg)	2.9	1.9, 4.5	2.3	1.9, 2.9	0.49*	0.48*	0.50*	1.2	0.83–1.8
Vitamin B12 (μg)	4.4	2.2, 9	3.9	2.6, 6.1	0.10***	0.01***	–0.01***	1.1	0.39–3.2
Folate (μg)	234	125, 441	152	97, 239	0.53*	0.49**	0.51*	1.5	0.71–3.3
Vitamin C (mg)	63	23, 168	39	17, 88	0.47*	0.51*	0.53*	1.6	0.37–6.9

FFQ, food frequency questionnaire; 24HDR, 24-hour dietary recall. *$P < 0.01$(2-tailed); **$P < 0.05$ (2-tailed); ***$P > 0.05$ (2-tailed). †log transformed nutrient intakes ($FFQ_{log} - 24HDR_{log}$).

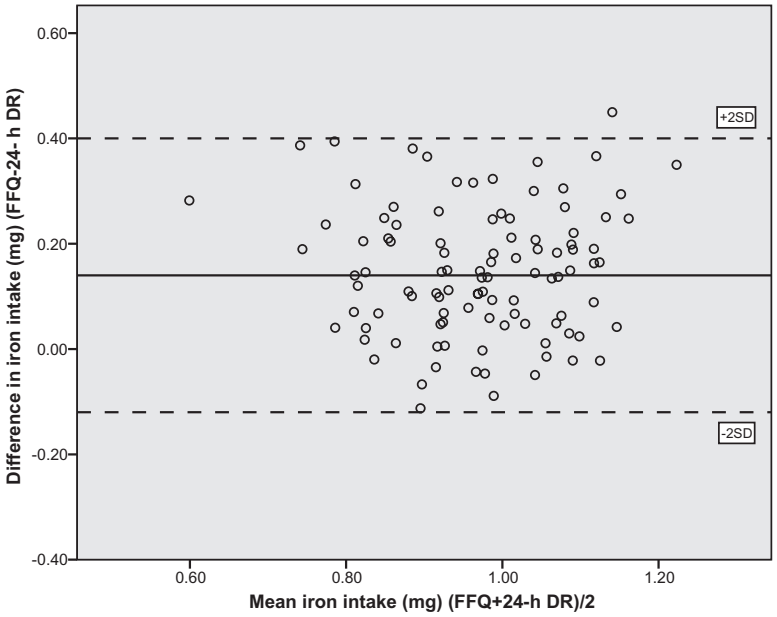

Fig. 2. Bland-Altman plot between the FFQ and the 24HDRs for log-transformed crude iron intakes (mg).
Solid line – mean difference; dashed lines – plus or minus 2 standard deviations (SD) ratio.

intakes and refers to crude log-transformed intakes. All mean differences were positive i.e. higher intakes were reported by the FFQ than the 24HDRs with the exception of vitamin C (Table 3). Average differences expressed by the mean ratio suggested FFQ overestimations by 50% for selenium, folate, vitamin C and vitamin D. The lower and upper LoA represented the range in which 95% of the differences between the two methods were expected to lie. Widest ranges of LoA were observed for vitamin C, selenium and vitamin D.

Agreement for quintiles was calculated for crude and energy-adjusted nutrient intakes estimated by the FFQ1 and the 24HDRs (Table 4). Percentage of individuals classified into the same quintile varied from 55% for vitamin B_{12} to 77% for magnesium. An average of 66% of nutrient intakes was assigned to the same quintile by the two methods. Following adjustments, the proportion of individuals classified into the same quintile improved only by 2%. Improvement was observed for 11 out of the 24 nutrient examined. Folate was least and magnesium was most affected (range 1% to 11% improvement). The percentage of individuals classified into the extreme quintile of the distribution varied from 0% for magnesium to 6.7% for vitamin D and an overall misclassification of 3% was observed. Energy-adjustments had an insignificant effect on the percentage misclassification. Due to the high number of participants reporting not consuming alcohol, it was not included in the analysis.

Discussion

This research was undertaken as part of an on-going cohort study aiming to provide information on the dietary behaviour of postpartum women living in the North of the UK and their relationship to infant feeding practices. To the authors' knowledge, this is the first reproducibility and validity study of an FFQ carried out on a postpartum population in the UK, which is an important step towards increasing awareness of the importance of the nutritional status of such vulnerable populations. We found good sized correlations for nutrient intakes assessed by the FFQs completed in one month and fair correlations for the majority of nutrient intakes assessed by the FFQ and the mean of the 24HDRs.

Mean intakes changed slightly between the two FFQ administrations. Our reproducibility coefficients compare well with those of other studies which usually report values ranging from 0.50 to 0.70 (Cade

Table 4. Cross-classification of nutrient intakes quintiles from the FFQs and the 24HDRs

Nutrient	Absolute intake		Energy-adjusted intake*	
	Correctly classified (%)	Grossly misclassified (%)	Correctly classified (%)	Grossly misclassified (%)
Energy (MJ)	64	3.8		
Protein (g)	64	2.9	64	2.9
Total fat (g)	70	4.8	67	5.8
Carbohydrate (g)	60	3.8	68	4.8
Saturated fatty acids (g)	64	2.9	64	2.9
Monounsaturated fatty acids (g)	62	5.8	65	6.7
Polyunsaturated fatty acids (g)	66	0.1	64	4.8
Sugars (g)	63	2.9	67	4.8
Starch (g)	62	3.8	54	4.8
Dietary fibre (g)	66	1	66	1.9
Calcium (mg)	59	3.8	64	5.5
Magnesium (mg)	67	1	78	1
Iron (mg)	75	1	81	1.9
Zinc (mg)	62	0.1	65	3.8
Manganese (mg)	77	0	84	1
Selenium (μg)	65	4.8	64	5.8
Vitamin D (μg)	64	6.7	63	5.8
Vitamin E (mg)	63	5.8	59	1.9
Thiamine (mg)	66	1.9	72	2.9
Riboflavin (mg)	74	3.8	68	1.9
Vitamin B6 (mg)	70	3.8	77	1.9
Vitamin B12 (μg)	55	1	54	4.8
Folate (μg)	74	1	75	1.9
Vitamin C (mg)	74	1.9	73	1.9

FFQ, food frequency questionnaire; 24HDR, 24-hour dietary recall. *Energy-adjusted prior to quintile classification.

et al. 2002). Recently conducted studies have reported mean values of 0.71 in the California Teachers Study (Horn-Ross *et al*. 2008), 0.46 in a sample of Chinese pregnant women (Cheng *et al*. 2008) and coefficients values varied from 0.39 to 0.83 in a general sample of French participants (Deschamps *et al*. 2007). Studies conducted in US and Finnish pregnant populations have also reported reasonably high correlations (Suitor *et al*. 1989; Errkola *et al*. 2001). Mean correlation following energy adjustments was slightly reduced. Coefficient reductions affected most macronutrients and to a lesser extent the micronutrients examined. A substantial reduction was observed for total fat, carbohydrate and thiamine. Reduction of reproducibility correlations among others for total fat and carbohydrates attributed to a decrease of the between-subject variability in nutrient intakes following controlling for energy intake has also been described in other studies (Bohlscheid-Thomas *et al*. 1997; Nagel *et al*. 2007).

As expected, the FFQs provided higher estimates for all nutrients examined compared with the 24HDRs-except for vitamin E. Mean nutrient estimates assessed by the FFQs were around 30% higher than those estimated by the 24HDRs. Similar overestimates have also been reported in other validation studies when intakes assessed by FFQs were compared with those assessed by 24HDRs or food records (Suitor *et al*. 1989; Robinson *et al*. 1996; Errkola *et al*. 2001; Baer *et al*. 2005; Boucher *et al*. 2006).

The choice of the 24HDR as a reference method was based on the assumption that the response rate will be higher compared with other more precise but more subject-burdensome methods e.g. weighted food records. The nature of the reference method itself is one of the methodological factors that have an influence on the validity of the FFQ. Taking into account that recalls often underestimate intakes, and that the FFQ covers a longer time span, relatively low energy intakes obtained by the recall method, might

possibly indicate that mean FFQ energy intake was closer to the true value. Differences between the two methods might also reflect over reporting of the food items included in the FFQs or under reporting of frequency and quantity in the 24HDRs of foods consumed or to a certain extent both. Differences might have also been resulted due to the use of standard portion sizes in the questionnaire. This is likely to have led to under- or over-representation of intakes of those women who consumed relatively smaller or larger portions of foods and had difficulties accurately judging the frequency of consumption of specific food items.

The results of the Bland-Altman analysis indicated that mean differences were positive i.e. higher intakes were reported by the FFQ and wide LoA indicated misreporting of intakes for the majority of nutrients. The plots suggested no dependency between difference and mean value, suggesting that agreement between the FFQ and the 24HDRs was of the same magnitude irrespective of mean quantity. Only for few nutrients, including vitamin C, did the plot indicate increasing negative differences with increasing mean quantity, indicating that compared with the FFQ, the 24HDR underestimated intakes (data not shown). Similar trend have also been reported in other studies, where a trend in bias reflected by a tendency for the mean difference to rise or fall with increasing magnitude was observed (Robinson *et al*. 1999; Andersen *et al*. 2004; Brantsæter *et al*. 2008).

In this study, mean unadjusted correlation coefficient was 0.34 indicating fair agreement and values are in line with findings of other published studies (Cade *et al*. 2002). The findings suggest that the FFQ performed well across a range of nutrients and no particular pattern was observed. The highest unadjusted correlations were observed for manganese, iron, folate and thiamine, and lowest for vitamin B_{12} and vitamin E. We are aware of two other studies examining the validity of an FFQ to measure dietary intakes of US lactating women. Both studies, however, reported very poor agreement between the FFQ and 7-day food records and 24HDRs, respectively (Stuff *et al*. 1983; Forsythe & Gage 1994). The findings of a recent validation study conducted on a sample of 119 Norwegian pregnant women reported a similar average value of 0.36 (Brantsæter *et al*. 2008). The authors concluded that the level of agreement between the FFQ and the 4-day weighted food diary was satisfactory, and that when compared with biological markers, the FFQ was able to distinguish between low and high intakes.

Energy-adjustments failed to strengthen observed correlations, as values for the majority of nutrients were not appreciably affected. Energy-adjustment usually increases the correlations in cases where variability of nutrient intake is related to energy intake and decreases the correlation when variability is related to under- and overestimation (Willett 1998). In addition, corrections for within-person variability in the 24HDRs had no effect on the mean value, suggesting that collection of more than one observation per participant improved the true estimates. Martin-Moreno *et al*. (1993) reported an insignificant effect of energy-adjustment on correlation values, whereas Slater *et al*. (2003) observed decrease in values following adjustments. Katsouyanni *et al*. (1997), on the other hand, reported simultaneous increases and decreases in correlation values for the different nutrients analysed among Greek school teachers.

For some nutrients, lower correlations observed in our sample compared with other studies could be attributed to some of our population characteristics known to have an influence on the strength of the association between measurement methods (Gibney *et al*. 2004). In addition, findings between validation studies might not necessarily be directly comparable as the result of differences in the FFQ length, sample size, population characteristics, number of nutrients assessed, use of reference method and number of recording days. Even in studies where an FFQ was applied in the same population but its validity examined in a wider range of nutrients, differences in correlation values were reported (Suitor *et al*. 1989; Wei *et al*. 1999). This was also observed in our study when results were related to those of a study conducted in a group of pregnant women living in Sheffield, UK, with similar characteristics (Mouratidou *et al*. 2006a). In both instances, the FFQ provided higher energy intakes compared with the reference 24HDR method, but the FFQ in the postpartum study

performed better for the majority of nutrients. Differences in the number of recording days and physiological state of study participants could explain some of the observed differences.

Several participants' characteristics could impact on the outcome of a validation study including age, BMI, and health status in addition to the type of diet consumed; therefore, it is important that the study sample is similar to the main study population (Cade *et al*. 2002). Participants excluded from our validation study and those from the general HS sample had similar non-dietary characteristics given the strict eligibility criteria followed in the recruitment process. Taking into account observed similarities in their general characteristics, it is highly unlikely that their diets will differ as evidence suggest that the diets of women and/or groups from low-socioeconomic backgrounds are similarly unvaried and of low dietary quality (Lawrence *et al*. 2008; Lawrence & Barker 2009). It is also, highly unlikely that season of assessment affected participants' classification given that the 62-item FFQ did not ask many season specific food questions, in addition to lack of dietary diversity observed in low-income groups (Fowke *et al*. 2009; Lawrence & Barker 2009).

In spite of fair correlations for the majority of nutrients examined, the FFQ was able to distinguish between high and low consumers for all nutrients under investigation. The results of cross-classification showed that extreme misclassification of nutrient intake was rare. This suggests that the FFQ can appropriately rank individuals relative to one another even if absolute intakes may not be precise. Percentage agreement and misclassification in our study are very similar or slightly higher to those reported elsewhere (Errkola *et al*. 2001; Baer *et al*. 2005; Brantsæter *et al*. 2008). Two thirds of our participants were assigned to the same quintile by the two methods and around 3% in the extreme opposite. Despite this, a considerable amount (nearly 30%) of participants were misclassified into different categories of consumption. It is possible that this incorrect classification reflects not only some of the criticisms related to the test measure and its ability to assess absolute intakes but also some of the limitations of the reference method used discussed earlier (Willett 1998; Gibney *et al*. 2004).

The present analysis suggested that the FFQ applied in the HS study performed well in estimating intakes of a number of nutrients when compared with intake estimates obtained by repeated recalls. In addition, the findings suggested that the FFQ might under- and overestimate intakes for some nutrients. These results need to be taken into consideration when interpreting the results of the HS study. In conclusion, the results of the present study suggested fair reproducibility and validity of the majority of nutrients examined as assessed by an interviewer-administered FFQ and the reference method. The study findings suggest that the FFQ is a valid tool for ranking individuals according to nutrient distribution.

Acknowledgements

We would like to thank all the participants in the study. Jean M Russell MA MSc FSS, Computer Officer/Statistician, University of Sheffield. Sarah Wademan, Helen Nicholson, Rebecca Roberts, Jo Pearce, Antony Renshaw, Kate Percival, Stephanie Parker, Fazeelat Akhtar, Catherine Cliff and Delia Kenley nutrition researchers at the University of Sheffield.

Sources of funding

The research on which this paper is based was funded by a programme grant from The Leverhulme Trust (award number F/00118/AQ).

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

Andersen I.F., Lande B., Trygg K. & Hay G. (2004) Validation of a semi-quantitative food frequency questionnaire used among 2-year-old Norwegian children. *Public Health Nutrition* **7**, 754–764.

Baer H.J., Blum R.E., Rockett H.R., Leppert J., Gardner J.D., Suitor C.W. *et al.* (2005) Use of a food frequency questionnaire in American Indian and Caucasian pregnant women: a validation study. *BMC Public Health* **15**, 135.

Birch L.L. & Fisher J.O. (1998) Development of eating behaviours among children and adolescents. *Pediatrics* **101**, 539–549.

Bland J.M. & Altman D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* **1**, 307–310.

Bohlscheid-Thomas S., Hoting I., Boeing H. & Wahrendorf J. (1997) Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. *European Prospective Investigation into Cancer and Nutrition International Journal of Epidemiology* **26**, S59–S70.

Boucher B., Cotterchio M., Kreiger N., Nadalin V., Block T. & Block G. (2006) Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. *Public Health Nutrition* **9**, 84–93.

Brantsæter L.A., Haugen M., Alexander J. & Meltzer H.M. (2008) Validity of a new food frequency questionnaire for pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). *Maternal and Child Nutrition* **4**, 28–43.

Cade J., Thompson R., Burley V. & Warm D. (2002) Development, validation and utilization of food frequency questionnaires – a review. *Public Health Nutrition* **5**, 567–587.

Cheng Y., Yan H., Dibley M.J., Shen Y., Li Q. & Zeng L. (2008) Validity and reproducibility of a semi-quantitative food frequency questionnaire for use among pregnant women in rural China. *Asia Pacific Journal of Clinical Nutrition* **17**, 166–177.

Department of Health (2002) *Healthy Start: Proposals for reform of the Welfare Food Scheme*. Department of Health: London.

Deschamps V., de Lauzon-Guillain B., Lafay L., Borys J.M., Charles M.A. & Romon M. (2007) Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. *European Journal of Clinical Nutrition* **63**, 282–291.

Doran L. & Evers S. (1997) Energy and nutrient inadequacies in the diets of low-income women who breast-feed. *Journal of the American Dietetic Association* **97**, 1283–1287.

Errkola M., Karppinen M., Javanainen J., Rasanen L., Knip M. & Virtanen S.M. (2001) Validity and reproducibility of a food frequency questionnaire for pregnant Finnish women. *American Journal of Epidemiology* **154**, 466–476.

Ford F.A., Mouratidou T., Wademan S.E. & Fraser R.B. (2009) Effect of the introduction of 'Healthy Start' on dietary behaviour during and after pregnancy: early results from the 'before and after' Sheffield study. *British Journal of Nutrition* **101**, 1828–1836.

Forsythe H.E. & Gage B. (1994) Use of a multicultural food-frequency questionnaire with pregnant and lactating women. *American Journal of Clinical Nutrition* **59**, 203S–206S.

Fowke J.H., Schlundt D., Gong Y., Jin F., Shu X.O., Wen W. *et al.* (2009) Impact of season of food frequency questionnaire administration on dietary reporting. *Annals of Epidemiology* **14**, 778–785.

George G.C., Hanss-Nuss H., Milani T.J. & Freeland-Graves J.H. (2005) Food choices of low-income women during pregnancy and postpartum. *Journal of the American Dietetic Association* **105**, 899–907.

Gibney M.J., Margetts B.M., Kearney J.M. & Arab L. (2004) *Public Health Nutrition*, 1st edn, Blackwell Publishing: London.

Horn-Ross P.L., Lee V.S., Collins C.N., Stewart S.L., Canchola A.J., Lee M.M. *et al.* (2008) Dietary assessment in the California Teachers Study: reproducibility and validity. *Cancer Causes and Control* **19**, 595–603.

Jain A., Sherman S.N., Chamberlin L.A., Carter Y., Powers S.W. & Whitaker R.C. (2001) Why don't low-income mothers worry about their preschoolers being overweight? *Pediatrics* **107**, 1138–1146.

Katsouyanni K., Rimm E.B., Gnardellis C., Trichopoulos D., Polychronopoulos E. & Trichopoulou A. (1997) Reproducibility and relative validity of an extensive semi-quantitative food frequency questionnaire using dietary records and biochemical markers among Greek schoolteachers. *International Journal of Epidemiology* **26**, 118–127.

Lawrence W. & Barker M. (2009) A review of factors affecting the food choices of disadvantaged women. *Proceedings of the Nutrition Society of* **2**, 189–194.

Lawrence W.T., Skinner T.C., Haslam C.O., Robinson S.M., Inskip H.M., Margetts B.M. *et al.* (2008) Constraints on food choices of women in the UK with lower educational attainment. *Public Health Nutrition* **12**, 1229–1237.

Mackey A.D., Picciano M.F., Mitchell H. & Smiciklas-Wright H. (1998) Self-selected diets of lactating women often fail to meet dietary recommendations. *Journal of the American Dietetic Association* **98**, 297–302.

Martin-Moreno J.M., Boyle P., Gorgojo L., Maisonneuve P., Fernandez-Rodriguez J.C., Salvini S. *et al.* (1993) Development and validation of a food frequency questionnaire in Spain. *International Journal of Epidemiology* **22**, 512–519.

Mouratidou T., Ford F. & Fraser R.B. (2006a) Validation of a food-frequency questionnaire for use in pregnancy. *Public Health Nutrition* **9**, 515–522.

Mouratidou T., Ford F. & Fraser R.B. (2006b) Dietary assessment of a population of pregnant women in Sheffield UK. *British Journal of Nutrition* **96**, 929–935.

Nagel G., Zoller D., Ruf T., Rohrmann S. & Linseisen J. (2007) Long-term reproducibility of a food-frequency questionnaire and dietary changes in the European Pro-

spective Investigation into Cancer and Nutrition (EPIC)-Heidelberg cohort. *British Journal of Nutrition* **98**, 194–200.

Noble M., Wright G. & Dibben C. (2004) *Indices of Deprivation 2004. Report to the Office of the Deputy Prime Minister*. Neighbourhood Renewal Unit: London.

Rees G.A., Doyle W., Srivastava A., Brooke Z.M., Crawford M.A. & Costeloe K.L. (2005) The nutrient intakes of mothers of low birth weight babies – a comparison of ethnic groups in East London, UK. *Maternal and Child Nutrition* **1**, 91–99.

Robinson S., Skelton R., Barker M. & Wilman C. (1999) Assessing the diet of adolescent girls in the UK. *Public Health Nutrition* **2**, 571–577.

Robinson S.M., Godfrey K., Osmond C., Cox V. & Barker D. (1996) Evaluation of a food frequency questionnaire used to assess nutrient intakes in pregnant women. *European Journal of Clinical Nutrition* **50**, 302–308.

Rogers I., Emmett P. & the ALSPAC Study T. (1998) Diet during pregnancy in a population of pregnant women in the South West England. *European Journal of Clinical Nutrition* **52**, 246–250.

Slater B., Philippi S.T., Fisberg R.M. & Latorre M.R. (2003) Validation of a semi-quantitative adolescent food frequency questionnaire applied at a public school in São Paulo, Brazil. *European Journal of Clinical Nutrition* **57**, 629–635.

Stuff J.E., Garza C., Smith E.O., Nichols B.L. & Montandon C.M. (1983) A comparison of dietary methods in nutritional studies. *American Journal of Clinical Nutrition* **37**, 300–306.

Suitor C.J., Gardner J. & Willett W. (1989) A comparison of food frequency and diet recall methods in studies of nutrient intake of low-income pregnant women. *Journal of the American Dietetic Association* **89**, 1786–1794.

Wei E.K., Gardner J., Field A.E., Rosner B.A., Colditz G.A. & Suitor C.W. (1999) Validity of a food frequency questionnaire in assessing nutrient intakes of low-income pregnant women. *Maternal and Child Health Journal* **3**, 241–246.

Willett W., Howe G.R. & Kushi L.H. (1997) Adjustment for total energy intake in epidemiological studies. *American Journal of Clinical Nutrition* **65**, 1220S–1228S.

Willett W.C. (1998) *Nutritional Epidemiology*, 2nd edn. Oxford University Press: New York.

World Health Organization, Regional Office for Europe (2007) *European Action Plan for Food and Nutrition Policy 2007–2012*. Available at: http://www.euro.who.int/Document/E91153.pdf